Research Paper

International Security Programme

June 2024

The role of the private sector in combatting gendered cyber harms

How private sector technologies can both endanger and advance gender-transformative cybersecurity

Isabella Wilkinson, Julia-Silvana Hofstetter,
James Shires and Mardiya Siba Yahaya

CHATHAM
HOUSE

Chatham House, the Royal Institute of International Affairs, is a world-leading policy institute based in London. Our mission is to help governments and societies build a sustainably secure, prosperous and just world.

Contents

Summary

- This research paper argues that private sector actors play a crucial part in generating, enabling and mitigating gendered cyber harms. They can both seriously endanger and meaningfully advance gender-transformative cybersecurity – cybersecurity defined broadly to challenge harmful gender norms, roles and realities, and to consider a wide range of gendered harms experienced in cyberspace.

- The private sector has a crucial role to play in advancing gender-transformative cybersecurity. As developers, designers and deployers of technology, they bear significant responsibility for ensuring that their technologies do not lead to, or exacerbate, gendered cyber harms. Given these important responsibilities, private sector actors should commit to recognizing and addressing such harms and advancing a stable, secure cyberspace for all.

- Technology design choices have implications for security at all levels: from an individual's experience of cybersecurity to the implications of technology for national and international security. Features of technology design can be commercialized and 'weaponized' by third parties. However, technologies developed by the private sector can also be redesigned and readapted for new contexts, including those that seek directly to counter or mitigate gendered cyber harms.

- Design and redesign choices about technologies and the protections afforded to users do not take place in a vacuum. They are influenced by a complex array of internal and external pressures on those responsible for developing and deploying those technologies, ranging from gendered assumptions to public outcry and criticism. Choices made in this process can contribute to gendered cyber harms by incorporating explicit and implicit gendered structures, ideals and assumptions, which are then propagated and reinforced as technologies are adopted – even if inadvertently.

- Chapters 1 and 2 of this paper consider the complex interactions between gender, technology design, the private sector and security. The paper then presents a series of case studies showing the real-world effects of gendered cyber harms. Chapter 3 documents cyber harms experienced by queer social media users in Nigeria and South Africa, while Chapter 4 exposes how harms can be generated through the commercialization and 'weaponization' of sensitive data – with reference to examples from the US and the Middle East and North Africa. Chapter 5 shows how technologies can be used to improve cybersecurity and combat gender-based violence, citing examples from South Korea and India.

- The paper then concludes with a set of recommendations for private sector stakeholders seeking to design and deploy secure and gender-transformative technologies. Specific measures include:
  - Critically evaluating data sharing and cooperation with state entities, adopting a human rights and gender perspective to map potential harms;
  - Assessing the efficacy of user privacy and data sharing settings for all technologies, and the accessibility and ease of changing those options;
  - Mapping technological relationships with commercial partners and potential risks;
  - Implementing additional technical features and mitigations that enable users to reduce risks;
  - Incorporating user experiences and feedback into technology design, redesign and readaptation; and
  - Building independent internal gender expertise and connecting to international networks seeking to establish and promote best practices.

01 Introduction

While states must enable the right environment, the private sector is responsible for ensuring that gendered dimensions are accounted for in their design, development and deployment of technologies.

Gender-transformative cybersecurity is cybersecurity defined broadly to challenge harmful gender norms, roles and realities, and to consider a wide range of gendered harms experienced in cyberspace. Its aim is to build a cyberspace that does not spread, amplify or replicate gendered inequalities and insecurities that are prevalent offline.[1]

Advancing gender-transformative cybersecurity requires the involvement of multiple stakeholders. States and public sector actors bear responsibility for enabling and encouraging governance that elevates human security – for people of all genders – to the same level as national or international security. Private sector actors (including technology developers) also hold significant responsibilities in ensuring that technology design, development and deployment account for gendered impacts and that they help to advance gender equality – particularly given that technologies can result in gendered harms.

This research paper investigates the role of private sector entities in gendered cyber harms. It asks: **how do technologies designed by the private sector contribute to and/or mitigate gendered cyber harms?** Building on a previous Chatham House publication,[2] this paper argues that technologies developed by the private sector contribute to gendered cyber harms by incorporating explicit and implicit gendered structures, ideals and assumptions, which are then propagated and reinforced as technologies are adopted.

[1] Shires, J., Hassib, B. and Swali, A. (2024), *Gendered hate speech, data breach and state overreach: Identifying the connections between gendered cyber harms to shape better policy responses*, Research Paper, London: Royal Institute of International Affairs, https://doi.org/10.55317/9781784135973.
[2] Ibid.

Technologies designed by the private sector may have the potential for both enabling and mitigating harm. External evaluation and criticism can result in the redesign of these technologies and their readaptation for new uses and in new contexts, including for directly seeking to address gendered impacts.

The paper is structured as follows. First, it considers the complex interactions between gender, technology, technology design, the private sector and security. It then presents empirical case studies showing the real-world effects of gendered cyber harms, before concluding with a set of general recommendations intended to assist private sector stakeholders seeking to design and deploy secure and gender-transformative technologies.

02 Private sector technologies and gendered cyber harms

Gender and technology design are interrelated, with technology design incorporating and influencing gender both explicitly and implicitly.

This chapter introduces three premises to help understand the role of the private sector in both contributing to and mitigating gendered cyber harms. First, gender and technology design are interrelated, with technology design incorporating and influencing gender both explicitly and implicitly.

Second, companies in the private sector are technology designers, implementers and market shapers, primarily motivated by profit. Understanding the strength of, and limits to, the profit motive is essential for understanding private sector decision-making and the ways in which design choices lead to, or mitigate, gendered cyber harms.

Third, private actors can also respond to criticism through redesign. Technologies intended for use in one context can be readapted for use in different contexts by other actors (for instance, if a certain technology is used as a monitoring mechanism by law enforcement in their investigations).

2.1 Gender, technology design and international (cyber)security

In general, technologies are developed to solve problems, including security problems. However, the development of technology can overlook security considerations, and their deployment can cause or exacerbate security issues. Various social influences drive the emergence of different technologies, reflecting assumptions about what problems need to be solved and whose security considerations need to be prioritized and included in solutions. Gender is one of the social structures that has an influence.

Technology design can incorporate gender in two ways: explicitly and implicitly. Explicit gendering occurs when actors consciously incorporate gender into technology design: whether by broadly designating products for use by exclusively men or women, or through more subtle, ideal/typical notions of masculinity and femininity. Modern technologies, such as cars or home appliances, import extensive gendered connotations and associations due to their emergence in specific social contexts, where they were used for gendered roles. They also import gendered connotations and associations due to the deliberate choices of marketing and advertising campaigns, which have often sought to create an explicitly gendered image of those technologies.

Explicit and implicit gendering can occur together. Indeed, nearly all technologies are likely to incorporate both implicit and explicit gendered elements.

The second way in which technology design incorporates gender roles and identities is implicit. It is the idea of a gendered default, where technologies are designed for a particular physical body or set of psychological characteristics or needs.[3] Often, this default is expressed by the assumption of heterosexual masculinity. The creation of virtual reality headsets too heavy for the average woman,[4] or mobile phones too big to fit comfortably in smaller hands, is evidence of this kind of implicit gendering. In such cases, a likely unstated technology design assumption was that such devices would be used by men. This could be the result of technologies designed, developed and tested by teams with poor gender diversity, or with a low awareness of the gendered implications of their design choices. Implicit gendering of technologies is not, however, only oriented towards the needs of men: for example,

[3] In academic circles, this is often known as 'androcentrism': the practice of centralizing a masculine perspective and worldview. For further reading, see Criado-Perez, C. (2020), *Invisible Women: Exposing Data Bias in a World Designed for Men*, London: Chatto & Windus.
[4] McEvoy, C. (2023), 'The world is built for men. Is virtual reality next?', Culture 3 blog, 12 December 2023, https://www.culture3.com/posts/the-world-is-sexist-built-for-men-is-virtual-reality-next; Stanney, K., Fidopiastis, C. and Foster, L. (2020), 'Virtual Reality Is Sexist: But It Does Not Have To Be', *Front. Robot. AI Sec. Virtual Environments*, 7(2020), pp. 1–19, https://doi.org/10.3389/frobt.2020.00004.

the size, portability and design of typewriters changed along with their gendered connotations.[5] Explicit and implicit gendering can occur together. Indeed, nearly all technologies are likely to incorporate both implicit and explicit gendered elements.

The relationship between technology and gender is bi-directional. This means that, because gender influences technology design, technology design choices can also help to establish and revise concepts of gender. An example may be the algorithmic design of social media platforms enabling – or jeopardizing – space for discussing expression of gender identity. Because neither technology design nor gender are static (i.e. both may change over time), their mutual relationship can create complex feedback loops in which the design of the technology changes along with its gendered connotations.

Complex interactions between technology and gender have implications for international security. As detailed by feminist scholars of international relations and security, conventional approaches to security – both in scholarship and in practice – are predominantly masculinist.[6] Such approaches are exclusionary and overlook gendered implications (for instance, by overlooking the gendered differences in experiences of being and feeling 'secure'). They may also overlook the way in which technology reflects, embeds and instrumentalizes power, and the implications this carries for human security.

Digital technologies have well-established implications for international security. A topical example may be the use of artificial intelligence (AI)-generated media to advance threats (such as deepfakes and disinformation targeting politicians). These implications are gendered: the use of deepfakes for perpetrating image-based abuse has been well documented,[7] and women in politics are far more likely to experience social media abuse than their male counterparts.[8] It is essential to consider both the 'traditional' and human security implications of the gendered use and misuse of technologies.

Private actors can operate and deploy technologies on a global scale. They are involved in the provision of cybersecurity solutions at multiple levels, and thus involved in and crucial to the development and maintenance of a stable, secure cyberspace. As Section 2.2 explores, although the relationship between cyber insecurity, gender and technology design is experienced on an individual level, it can be manifested structurally as a security challenge faced at the community and group level, and even globally.

[5] Plotnick, R. (2019), 'Tethered women, mobile men: Gendered mobilities of typewriting', *Mobile Media & Communication*, 8(2), pp. 188–208, https://doi.org/10.1177/2050157919855756.
[6] Tickner, J. A. (2004), 'Feminist responses to international security studies', *International Peace Review*, 16(1), pp. 43–48, https://doi.org/10.1080/1040265042000210148.
[7] Davidson, J. et al. (2019), *Adult Online Hate, Harassment and Abuse: A Rapid Evidence Assessment*, report, University of East London, UK Council for Internet Safety and London School of Economics and Political Science, June 2019, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/811450/Adult_Online_Harms_Report_2019.pdf.
[8] See, for example, Amnesty International UK (2020), 'India: Women politicians face shocking scale of abuse on Twitter – new research', press release, 23 January 2020, https://www.amnesty.org.uk/press-releases/india-women-politicians-face-shocking-scale-abuse-twitter-new-research; Amnesty International UK (2020), 'UK: online abuse against black women MPs 'chilling'', press release, 9 June 2020, https://www.amnesty.org.uk/press-releases/uk-online-abuse-against-black-women-mps-chilling; Guerlin, C. and Maharasingam-Shah, E. (2020), *Public Figures, Public Rage: Candidate abuse on social media*, report, London: Institute for Strategic Dialogue, https://www.isdglobal.org/isd-publications/public-figures-public-rage-candidate-abuse-on-social-media.

Box 1. What are gendered cyber harms?

Gender is one of many social identities that impact how individuals perceive and experience cybersecurity. Cyber harms – and the vulnerabilities and risks that perpetuate, or amplify, these harms – thus differ based on an individual's gender identity.

There are three main kinds of gendered cyber harm:

- **Hate speech**, often via online harassment and abuse, and other content-based harms such as disinformation;
- **Data breach**, including privacy violations through hacking or leaking personal or sensitive data; and
- **State overreach**, such as cybercrime legislation reinforcing discriminatory gender norms.

These harms are cascading because one form of gendered cyber harm leads to another. They are compounding because such cascades increase the impact on the individual or individuals. In short, harms give rise to deeper harms. Understanding gendered cyber harms as cascading and compounding allows for a more holistic appreciation of how offline and online gender harms interact, intersect and reinforce one another.

Source: Shires, J., Hassib, B. and Swali, A. (2024), *Gendered hate speech, data breach and state overreach: Identifying the connections between gendered cyber harms to shape better policy responses*, Research Paper, London: Royal Institute of International Affairs, https://doi.org/10.55317/9781784135973.

2.2 The private sector, technology and gender

Technologies are designed, implemented and gendered in an ecosystem involving a range of different actors, including not just technology designers but also educators, policymakers, regulators and users. From technology design to deployment and evaluation, private sector entities have historically played a central role in importing gendered associations and connotations. Advancements in new technologies like generative AI threaten to amplify gendered harms both online and offline. This section describes the interaction of profitability, gender and technology design choices. It also outlines the internal and external pressures that result in the implicit and explicit gendering of technology, and in redesign.

Private sector entities are not only global leaders in technology development and deployment, but increasingly also in technology governance. Examining the private sector's role in technology design and gendered cyber harms is urgent for many reasons, one being that law enforcement agencies and public administration bodies around the world have rapidly adopted private sector-developed technology solutions. State adoption of private sector-developed technologies is particularly worthy of focus, as it presents an important example of where the encoding and amplification of harmful gendered connotations or bias could translate to offline and online harms.

The private sector is not a single set of actors, instead incorporating everything from venture capital firms to multinational technology companies; from small- and medium-sized businesses to state-owned enterprises. Any simplification of these different actors into a single category is bound to overlook relevant differences – for example, a publicly listed technology company that has published technology design principles to combat domestic abuse would likely respond differently to criticism of the gendered implications of its technologies than a privately owned social media platform. This paper predominantly focuses on technology companies designing and deploying their own technologies (with a particular focus on social media and messaging platforms and dating applications). It also analyses private sector entities that offer information services (like data brokers), and for the reasons noted above, state and police use (or, readaptation) of technology designed in the private sector.

Companies are motivated primarily by profit. As part of a broader social structure, gendered assumptions and roles are part of the external context in which they must manoeuvre to generate profit. More specifically, companies can also exploit gendered assumptions and differences to stand out against competitors. While this often leads to alignment with and reinforcement of gender stereotypes, such competition for distinction also motivates commercial decisions to subvert or challenge gendered norms and hierarchies. Of course, most private sector decisions do not deliberately seek to directly profit from, or perpetuate, gendered cyber harms. But even minor decisions may unwittingly enable harms, if bad choices (including those on design) can be exploited and 'weaponized' by third parties.

Technology companies play a unique role in the interaction between profitability, gender and technology. Like other private actors, technology companies make design choices to maximize profitability, the pursuit of which is conditioned by gendered assumptions. Some technology companies are both exceptionally profitable and exceptionally able to reach people across the world. The consequence of this is that, when faced with negative feedback (such as public pressure or punitive measures as a result of regulation), technology companies are in a complex position due to the financial implications of their perceived or actual ability to change the product. For many technology companies, one of their core concerns is to retain and grow their user base, including by maintaining trust and responding to feedback and competition (albeit to different and, in some cases, diminishing, degrees). This complex landscape is explored in the case studies.

The private sector's role in gendering technology design can be illustrated through the example of voice assistants, which both reflect and entrench gender stereotypes.[9] Companies producing voice-assistant systems initially tested

[9] West, M., Kraut, R. and Ei, C. H. (2019), *I'd blush if I could: closing gender divides in digital skills through education*, Paris: United Nations Educational, Scientific and Cultural Organization and EQUALS Skills Coalition, https://doi.org/10.54675/RAPC9356; Chin-Rothmann, C. and Robison, M. (2020), 'How AI bots and voice assistants reinforce gender bias', Brookings Institution, 23 November 2020, https://www.brookings.edu/articles/how-ai-bots-and-voice-assistants-reinforce-gender-bias; Steele, C. (2018), 'The Real Reason Voice Assistants Are Female (and Why It Matters)', PC Mag, 4 January 2018, https://uk.pcmag.com/smart-home/92697/the-real-reason-voice-assistants-are-female-and-why-it-matters; Abercrombie, G., Curry, A. C., Pandya, M. and Rieser, V. (2021), 'Alexa, Google, Siri: What Are Your Pronouns? Gender and Anthropomorphism in the Design and Perception of Conversational Assistants', *ARXIV*, https://arxiv.org/abs/2106.02578. Also, these assistants were initially used for satellite navigation, with the gendered connotations of map-reading adding a further layer to the gendered construction of such voices. See Fausto-Sterling, A. (2000), *Sexing the Body: Gender Politics and the Construction of Sexuality*, New York: Basic Books.

a range of voices (both masculine- and feminine-sounding), with commercial pressures incentivizing the design of systems where the voice was 'trusted' and not 'irritating' (both of which are highly gendered attributes).[10] In this way, technology developers in companies responded to external gendered pressures to produce a product that would be most well-received by customers and, therefore, both useful and profitable. The choice of a feminine-sounding voice perpetuates gender stereotypes and hierarchies by projecting assumed gendered roles in a professional environment into a digital device.

Once a technology has been released, redesign or rebranding decisions may be driven by reputation and profitability concerns, with companies responding to positive and/or negative feedback on their design choices. These pressures interact with others – both internal and external – such as: the interests of employees and stakeholders; abiding by best practices in corporate social responsibility; and acting in accordance with law, regulation and national and international norms.[11]

Unless criticism is particularly severe or such a decision is externally mandated, companies rarely respond by recalling or withdrawing their product and exiting the market. Sources of negative feedback range widely, including pressure generated by media reporting, what constitutes 'public opinion' or shareholder groups. As noted above, companies also make design and deployment decisions in response to national and international regulatory constraints (such as potential punitive measures) and norms (such as those relating to the responsible use of technology).[12]

An important subcategory of redesign emerges from technology design choices that directly or indirectly cause harm. Companies can act as inadvertent enablers of (gendered) harms, either due to third parties (ranging from law enforcement to criminal groups to other private sector actors) maliciously exploiting, weaponizing or otherwise misusing their products, or due to their own negligence in the design phase. An example is a geolocatable device, which presents security risks if abused. Risk emerges from the failure to consider and test for potential gendered harms during product design, potentially as a result of cost-saving measures (in other terms, the profit motive) or bias, but more likely because such testing is not deemed a priority or necessity. Further, companies may also directly profit from the use of their products to perpetuate gendered harms. In this case, they are not inadvertent enablers but rather facilitators. As explored in the case studies, the lines between inadvertent enabler and willing facilitator can be opaque and dynamic.

[10] Scott, S. (2007), 'Corporate Social Responsibility and the Fetter of Profitability', *Social Responsibility Journal*, 3(4), pp. 31–39, https://doi.org/10.1108/17471110710840215. For example, Apple no longer offered a default Siri voice in 2021, asking all iPhone users to choose, and offering a range of voices. However, whether gender-neutrality is possible has been questioned; see Hwang, G., Lee, J., Oh, C. Y. and Lee, J. (2019), 'It Sounds Like a Woman: Exploring Gender Stereotypes in South Korean Voice Assistants', *CHI Conference on Human Factors in Computing Systems*, May 2019, Paper No. LBW2413, pp. 1–6, https://doi.org/10.1145/3290607.3312915.
[11] For example, see Newman, L. H. (2023), 'This Clever New Idea Could Fix AirTag Stalking While Maximizing Privacy', WIRED, 27 December 2023, https://www.wired.com/story/apple-airtag-privacy-stalking-cryptographic-solution.
[12] Criticism or challenge occurs in, or on the edge of, prevalent gender structures, ideals and assumptions, as in some cases this criticism draws on hegemonic gender ideas. In other cases, the criticism is precisely of the reinforcement of those hegemonic ideas of gender.

Private developers of technology are increasingly aware of gendered cybersecurity risks.[13] Nonetheless, notable gaps in gender-sensitivity remain. In response to these gaps and the harms they perpetuate, a rich academic and practitioner movement has emerged against the 'male by default' approach to technology design. Practices they promote include participatory threat-modelling,[14] in which advocates for feminist technology design develop research and recommendations for developers of technology to identify and mitigate potential gendered cyber harms. Despite a promising upward trend in awareness, gendered security considerations do not yet play a substantial and sustained role in most private sector-led technology development, redesign and evaluation.

There are serious gaps in the design and implementation of gender-transformative technologies, which are compounded by the difficulties in incentivizing private actors to consider gender-transformative cybersecurity a profitable goal.

There are serious gaps in the design and implementation of gender-transformative technologies, which are compounded by the difficulties in incentivizing private actors to consider gender-transformative cybersecurity as a profitable goal. For the public sector, legislation, regulation, norms and best practices are important levers and can influence private sector activities. Governments wield an important, if not primary, influence in setting gender and technology policy agendas, which are responsive to gendered social structures and shape, to a degree, private sector activity.

In addition to the overarching role of law and regulation, technologies developed in the private sector can also be readapted for use by law enforcement or public sector actors.[15] Such readaptation can be organic or the result of formalized partnerships like public-private partnerships. A positive example explored in this paper is the readaptation of a social media messaging app for communicating with specialized teams in law enforcement to report potential crimes.[16] However, harms can also be exacerbated through this process of readaptation. An example given in this paper is the use of data from dating apps by law enforcement agencies

[13] For example, technology companies' trust and safety teams may scan the potential risk of abuse. See, for example, Meta (2024), 'New Tools to Help Protect Against Sextortion and Intimate Image Abuse', Meta blog, 11 April 2024, https://about.fb.com/news/2024/04/new-tools-to-help-protect-against-sextortion-and-intimate-image-abuse.

[14] For further reading, see Slupska, J., Duckworth, S., Ma, L. and Neff, G. (2021), 'Participatory Threat Modelling: Exploring Paths to Reconfigure Cybersecurity', *CHI Conference on Human Factors in Computing Systems*, May 2021, Article no. 329, pp. 1–6, https://doi.org/10.1145/3411763.3451731.

[15] This also ties into broader debates about what constitutes (digital) public infrastructure (DPI), as private sector actors are often the developers and implementers of DPI, such as payment, identity and data storage and exchange systems. For further discussion, see Eaves, D., Mazzucato, M. and Vasconcellos, B. (2024), 'Digital public infrastructure and public value: What is 'public' about DPI?' UCL Institute for Innovation and Public Purpose, Working Paper 2024/05, https://www.ucl.ac.uk/bartlett/public-purpose/publications/2024/mar/digital-public-infrastructure-and-public-value-what-public-about-dpi.

[16] See Chapter 5 and Jain, M. (2021), 'How WhatsApp became a tool for Indian police to fight harassment', Rest of World, 22 April 2021, https://restofworld.org/2021/how-whatsapp-became-a-tool-for-indian-police-to-fight-harassment.

in the Middle East and North Africa to target LGBTIQ+ individuals.[17] Some technologies developed by the private sector can be redesigned or readapted and deployed in a different context – for example, for mitigating gendered harms. However, without the right safeguards in its implementation, those technologies risk exacerbating existing harms.

2.3 Introducing the case studies

To investigate the roles and responsibilities of private sector actors in perpetuating and mitigating gendered cyber harms, this paper includes three chapters discussing real-world case studies: Chapter 3 considers social media use by individuals identifying as queer in Nigeria and South Africa; Chapter 4 considers the 'weaponization' of, reproductive health data in the US, and of LGBTIQ+ dating app data in the Middle East and North Africa; and finally, Chapter 5 considers detection technologies for images and videos of 'digital sex crimes' in South Korea, and online reporting mechanisms for gender-based violence in India.

These three chapters each encompass three interconnected kinds of gendered cyber harm – hate speech, data breach and state overreach. The case studies presented in each chapter span across cultural, national, and regional contexts. Chapter 3 is partially concerned with hate speech and other content-based harms, while Chapter 4 considers harms emerging from data breaches like privacy violations and state and police actors perpetuating harm. The two case studies in Chapter 5 address all three kinds of gendered cyber harm. As in the previous Chatham House paper on gendered cyber harms, these case studies advance the argument that gendered cyber harms are cascading and compounding. The case studies also emphasize the connections between offline and online harms, showing how cyber harms stem from and exacerbate offline prejudice, discrimination and violence.[18]

A central aim of the case studies is to explore how private sector actors can contribute to gendered cyber harms in different contexts. For this reason, the case studies occur at different points in the technology design and deployment cycle. Focusing mainly on social media and data markets, the first three case studies focus on how large-scale technology design – across whole sectors and markets – responds to both the profit motive and gendered social structures, resulting the commercialization and 'weaponization' of sensitive information. In contrast, the final two case studies focus on the readaptation and deployment of technology for monitoring and addressing gendered cyber harms, at the request of, or in tandem with, the public sector and law enforcement agencies.

These case studies foreground lived experiences, highlighting the basic point that, ultimately, it is people who interact with technologies and experience the harms they cause. Decisions on how and when to use the technologies in these case

[17] See Chapter 4 and Human Rights Watch (2023), *"All This Terror Because of a Photo": Digital Targeting and Its Offline Consequences for LGBT People in the Middle East and North Africa*, report, 21 February 2023, https://www.hrw.org/report/2023/02/21/all-terror-because-photo/digital-targeting-and-its-offline-consequences-lgbt.
[18] Shires, Hassib and Swali (2024), *Gendered hate speech, data breach and state overreach*.

studies represent nodes of criticism, challenge and even resistance. When someone restricts their location settings on a dating app, limits photo-sharing on social media, or chooses (not) to post an emergency message on an automated helpline, their decision is an expression of individual agency. It is therefore essential to foreground individual, human interactions with technology, and guard against an overly structural perspective.

03 The experiences of queer social media platform users

Gender-based targeting resulting from the insecure use of, and engagement with, social media platforms can lead to harms.

This chapter advances a definition of cybersecurity rooted in actions, practices, designs and policies, in addition to purely technical cybersecurity. It assesses the gendered cyber harms facilitated by visual media technologies and considers how a user's context (culture or location) may additionally shape these harms. Through six semi-structured interviews with purposefully sampled users from Nigeria and South Africa who self-identified as queer,[19] this case study demonstrates how large-scale technology design (across whole sectors and markets) has implications for gendered cyber harms and interacts with gendered social structures.

The two countries were chosen based on their advancement (South Africa) or lack thereof (Nigeria) of legislative protection for LGBTIQ+ people. While there are many similarities among the interviewees' experiences, this case study acknowledges the geographical, cultural and individual differences that exist between them. In doing so, it argues for placing people at the centre of the nexus of gender, technology and the private sector.

[19] 'Queer' is a general term adopted by some parts of the LGBTIQ+ community. Interviewees self-identified as queer, which is why this terminology is reflected in the title and text of this chapter. Participants for the interview were selected purposefully and supplemented using a snowballing technique. Six interviews were conducted in September 2023 and the conversations' duration was between one hour and one hour and 30 minutes via Zoom, and in-person with one participant. Three participants requested to stay anonymous. All interviews were recorded using a digital recorder. The interviewees were also provided a participant information sheet, and a consent form prior to the interviews. The data from the interviews were assessed largely through a thematic and narrative analysis.

3.1 Anxieties around escalation, scalability and visibility

An individual's context and identity affect their security needs, which are additionally correlated with *who* or *what* that individual trusts.[20] This point is captured by the concept of individuals having *differential* levels of vulnerability and trust, which carries gendered implications. All six interviewees expressed that they felt their cybersecurity vulnerabilities were shaped by a discriminatory social structure that determines their experiences of technology. From the perspective of four of the interviewees, it became clear that a key feature of TikTok's design – i.e. a reliance on extreme scalability and algorithmic escalation of content – can be a source of concern and depletes their trust in the platform.

Improvements to platform accountability and transparency can act as a form of gender-transformative cybersecurity.

On TikTok, a person's content or profile can go 'viral' without the creator's informed consent, and in ways that are not transparent to the creator. This may be due to algorithmic design choices like recommendation systems, which determine how content and profiles are scaled, prioritized or deprioritized by the platform. Metrics differ from platform to platform, but may include content length, caption, location and engagement. One interviewee from Nigeria, referred to as 'Seyi',[21] explained that the lack of transparency around TikTok's algorithm makes them unsure about how to present themselves on the internet. This is in addition to evolving threats and anxieties generated by and in online spaces. Their non-binary, feminist identity and positioning as a Nigerian on TikTok increases their fear of potential threats (e.g. from 'doxxing' and harassment, on a variety of platforms). These threats can be amplified through the scaling and virality of publicly produced content in ways that are unpredictable for users.[22]

Meanwhile, some interviewees from South Africa shared that their lack of trust in such platforms stems from witnessing identity-based threats and harms that others have experienced on TikTok, including death threats and stalking. As another interviewee, Arinze, said: 'Our collectivist nature mean we are in each other's business, and your content could end up on your aunt's 'For You' page.'

[20] Peña, P. (2023), *A Framework for Developing Gender-Responsive Cybersecurity Policy*, framework paper, Melville: Association for Progressive Communications, https://www.apc.org/sites/default/files/gender-cybersecurity-policy-norms.pdf.

[21] 'Seyi' is a pseudonym given by the authors to the participant, who requested to remain anonymous.

[22] In the case of TikTok, it commits to 'diversifying recommendations… to that end, sometimes you may come across a video in your feed that doesn't appear to be related to your expressed interests… This is an important and international component of our approach to recommendation: bringing a diversity of videos into your For You feed, giving you additional opportunities to stumble upon new content categories, discover new creators, and experience new perspectives and ideas as you scroll through your feed.' Although some external experts are invited to review TikTok's source code, the factors driving the personalization of user's For You pages are not public. Reportedly, they are based on factors like device settings, video information and user interactions. See TikTok (2020), 'How TikTok recommends videos #ForYou', TikTok Newsroom, 18 June 2020, https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you.

The false narrative of queer 'social contagion'[23] is used as a way to force queer people to live their online and offline lives in private. Content that has been scaled and amplified may generate anxieties about becoming the target of privacy violations as a result.[24] The harm of extreme scalability and algorithmic escalation is a locational dynamic, rooted in factors that dictate how technology users can express and perform their identity online, like politics, legislation and culture.

The queer users interviewed in Nigeria and South Africa perceive and experience gendered cyber harms posed by scalability and algorithmic escalation differently. 'Seyi' described TikTok as an extreme form of 'the public', which makes security difficult to navigate. On the other hand, Terrie, an interviewee from South Africa who is trans and identifies as a woman, felt she faced no imminent threat as a South African person using the platform to promote her YouTube content through a new medium. However, Terrie's experience was not shared by two other South African participants, who were more concerned with queer users receiving death threats.

The design of technologies and platforms creates an easily scalable and replicable form of threat because design choices make individuals who are already at-risk hyper-visible and vulnerable to greater and more diverse threats. Interviewees expressed concern with how threats become 'worse' when their content (unexpectedly) reaches thousands of people. Interviewees noted that while cyberspace poses a variety of risks, TikTok is particularly concerning because of how easy and quick it seems for content to 'automatically' escalate.[25] Scaling of content is a purposeful algorithmic design choice, often intended to maximize user engagement and, in turn, generate profit. In this case, unintended harms emerging from non-static design choices (i.e. regular adaptation of algorithms) are experienced differently by different users.

3.2 Navigating security, visibility and engagement

Responsibilities for ensuring security are shared among actors such as technology designers, regulators, enforcers and communities. With the right access to information and training, users can also make informed, communal choices to improve the safety and security of their own platform use. As some interviewees explained, specific engagement choices – such as choosing to not engage with or seek out harmful or abusive content – is a security choice, as it reduces the likelihood of being exposed to harmful content. However, if social media users – and particularly queer users – need to take additional steps to improve their online safety, technology companies also have a responsibility to transparently provide the necessary information and tools to enable this. Cybersecurity is embedded in sociotechnical dynamics and is influenced by one's gender, sexuality, race, class, location, religion and cultural context(s). The security needs of queer

[23] Ryan, H. (2023), 'Who's Afraid of Social Contagion?', Boston Review, 31 July 2023, https://www.bostonreview.net/articles/whos-afraid-of-social-contagion.
[24] Peña (2023), *A Framework for Developing Gender-Responsive Cybersecurity Policy*.
[25] See footnote 22 for further information on TikTok's recommendation system.

users reflect these dynamics in different ways, meaning (cyber)security is constantly co- or re-constituted in different locations and by different stakeholders. The potential for gendered cyber harms is difficult to measure, but technology companies and users can take steps both to prevent and mitigate risks.

Improvements to platform accountability and transparency can act as a form of gender-transformative cybersecurity. Platform accountability and transparency not only place the primary responsibility for ensuring safety and security on the provider (i.e. the platform), but also make it easier for users to make informed decisions on how to navigate or use these platforms with security in mind. Those responsible for designing, developing and regulating how platforms function, what information is shared with users, and how detailed that information is should integrate considerations of how straightforward it is for a user to action security measures (such as reporting and 'opt-out' functions). Such an approach would be gender-transformative because it recognizes, as many platforms already do, that gender-based targeting resulting from the insecure use of, and engagement with, social media platforms can lead to harm.

04 The 'weaponization' of geolocation data

Potential harms arise from technology design choices that enable third parties to commercialize, exploit and 'weaponize' data, leading to compounded harms when those data are used in gender-based targeting.

Technology results from, and is reliant on, a vast ecosystem of data-gathering. This system informs design choices and helps create more responsive, sophisticated and convenient technologies that, in turn, generate more data. Together with a growing market for third-party personal data, this cycle has led technology companies to maximize the collection of data, and to data becoming ever more profitable for technology companies and others. Consequently, commitments to ensuring data privacy for users have been less of a priority.

The risks of data privacy breaches are particularly severe for women and other marginalized groups, since discriminatory structures in society and in government institutions define whether and how personal data can be 'weaponized' to harm individuals and whether those data should, therefore, be considered 'sensitive'.[26] Thus, gender has an influence on the 'weaponizability' of seemingly non-sensitive data.

Potential harms arise from implicit technology design choices that enable third parties to commercialize, exploit and 'weaponize' data, leading to compounded harms when said data are used in gender-based targeting. The case studies in this chapter discuss gendered cyber harms associated with data privacy breaches in the context of commercially available data collected and aggregated by data brokers, and the 'weaponization' of geolocation data in particular. The chapter is divided into two case studies that consider two types of personal information

[26] Hofstetter, J-S. (2024), 'Gendered and Postcolonial Perspectives on Data Weaponization in Armed Conflict: The Case of Afghanistan', in Henshaw, A. and Mhajne, A. (eds), *Critical Perspectives on Cybersecurity: Feminist and Postcolonial Interventions*, Oxford: Oxford University Press, pp. 86–87 and pp. 96–99.

that are particularly sensitive from a gendered perspective. The first discusses the 'weaponization' of reproductive health-related data in the context of abortion criminalization in the US post-Roe vs Wade. The second discusses the potential for gendered cyber harms related to dating apps, and analyses the 'weaponization' of geolocation data obtained through LGBTIQ+ dating apps in the Middle East and North Africa. In both instances, cybersecurity has been of secondary importance to commercial incentives and priorities, creating an environment – and increasing the potential – of insecurity.

4.1 The role of data brokers

Data brokers aggregate a variety of data types from a vast range of sources to create highly detailed profiles of individuals.[27] Data brokers benefit from deliberate design choices that are intended to optimize and maximize data availability and collection. Mobile applications, social media platforms, search engines, website visits, online purchases, wearable devices and advertisement clicks are all data sources. Profiles obtained from these sources often include time-stamped geolocation data,[28] a type of data that is particularly problematic from a privacy perspective. While data brokers claim to protect individuals' data privacy rights by anonymizing their data, in many cases, geolocation data can be used to re-identify individuals and collate millions of data points aggregated in data brokers' profiles on those people.[29] Moreover, data managed by data brokers is at higher risk of being weaponized, since a wide range of actors (from law enforcement agencies to private individuals) can access the highly personal and sensitive data that brokers hold.

4.2 'Weaponization' of reproductive health data in the US

The process of seeking reproductive health information and services leaves a vast web of digital traces, whether through the use of digital devices and apps for fertility and period tracking; consulting telemedicine providers; purchasing healthcare products like abortion pills and pregnancy tests with a credit card; or simply searching for information online. While reproductive health-related information is a highly sensitive type of data, it is not always protected by high data-privacy standards and regulation. Much of this data can be collected and sold by data brokers.

[27] For further reading, see Sherman, J. (2021), *Data Brokers and Sensitive Data on U.S. Individuals*, Duke Sanford Cyber Policy Program, https://techpolicy.sanford.duke.edu/wp-content/uploads/sites/4/2021/08/Data-Brokers-and-Sensitive-Data-on-US-Individuals-Sherman-2021.pdf.

[28] Cyphers, B. (2022), 'How the Federal Government Buys Our Cell Phone Location Data', Electronic Frontier Foundation Deeplinks blog, 13 June 2022, https://www.eff.org/deeplinks/2022/06/how-federal-government-buys-our-cell-phone-location-data.

[29] Kagan, O. (2021), 'Can Location Data Truly Be Anonymized? New Research Says No', Fox Rothschild Privacy Compliance & Data Security blog, 30 March 2021, https://dataprivacy.foxrothschild.com/2021/03/articles/general-privacy-data-security-news-developments/can-location-data-truly-be-anonymized-new-research-says-no.

In recent years, digital services and products collecting health-related data have come under increased scrutiny[30] – most notably in the context of abortion criminalization in the US.[31] The overturning of the Roe v. Wade ruling by the US Supreme Court in 2022 ended the federal protection of abortion rights in the US and enabled abortion to be restricted in several US states.[32] Many people all over the US reacted by removing sensitive health apps and data from their digital devices due to concerns that this data could in future be obtained by law enforcement agencies and result in criminal prosecution.[33] In this context, most attention has been paid to the potential risks of data collected by fem-tech apps[34] and geolocation data tracking individuals' visits to abortion clinics.[35]

The design of fem-tech apps enables them to collect highly sensitive data not captured by other types of consumer devices.

There are currently more than 1,300 companies offering a range of fem-tech products aimed at helping women manage their health. These range from apps for mobile devices to Internet of Things products or worn devices.[36] The design of fem-tech apps enables them to collect highly sensitive data not captured by other types of consumer devices.[37] Here, users are assumed to be willing to divulge personal data in return for a service.

In 2021, the US Federal Trade Commission (FTC) issued a complaint against Flo Health, a fertility tracking app, for sharing sensitive health data from millions of individuals with marketing and analytics firms; the case was settled later in 2021, with the company being required to acquire users' 'affirmative consent

[30] Mehrnezhad, M., Van Der Merwe, T. and Catt, M. (2024), 'Mind the FemTech gap: regulation failings and exploitative systems', *Front. Internet Things*, 3(2024), pp. 1–14, https://doi.org/10.3389/friot.2024.1296599.
[31] Sherwood, H. (2024), 'Abortion prosecutions are never in the public interest, says royal college', *Guardian*, 22 January 2024, https://www.theguardian.com/world/2024/jan/22/illegal-abortions-prosecutions-uk-police-royal-college; Human Rights Watch (2023), 'Poland: Abortion Witch Hunt Targets Women, Doctors', 14 September 2023, https://www.hrw.org/news/2023/09/14/poland-abortion-witch-hunt-targets-women-doctors; Dellinger, J. and Pell, S. K. (2024), 'The criminalization of abortion and surveillance of women in a post-Dobbs world', Brookings Commentary, 18 April 2024, https://www.brookings.edu/articles/the-criminalization-of-abortion-and-surveillance-of-women-in-a-post-dobbs-world.
[32] *New York Times* (2024), 'Tracking Abortion Bans Across the Country', Updated 1 May 2024, https://www.nytimes.com/interactive/2022/us/abortion-laws-roe-v-wade.html.
[33] The decision to remove a healthcare app is simultaneously a reminder of the importance of individual choice in cybersecurity and that fem-tech devices and applications rely on user willingness to share their data. See Hill, K. (2023), 'Deleting Your Period Tracker Won't Protect You', *New York Times*, Updated 22 June 2023, https://www.nytimes.com/2022/06/30/technology/period-tracker-privacy-abortion.html.
[34] Logue, J. (2022), 'Users of "femtech" should be concerned–in a post-Dobbs world, their personal data could be used against them', Georgetown Law Technology Review, November 2022, https://georgetownlawtechreview.org/users-of-femtech-should-be-concerned-in-a-post-dobbs-world-their-personal-data-could-be-used-against-them/GLTR-11-2022.
[35] Cox, J. (2022), 'Data Broker Is Selling Location Data of People Who Visit Abortion Clinics', Vice Motherboard, 3 May 2022, https://www.vice.com/en/article/m7vzjb/location-data-abortion-clinics-safegraph-planned-parenthood; Grant, N. (2022), 'Google Says It Will Delete Location Data When Users Visit Abortion Clinics', *New York Times*, 1 July 2022, https://www.nytimes.com/2022/07/01/technology/google-abortion-location-data.html.
[36] FemTech Analytics (2021), 'FemTech Industry in Interactive Charts' https://www.femtech.health/interactive-charts.
[37] Mehrnezhad, M., Shipp, L., Almeida, T. and Toreini, E. (2022), 'Vision: Too Little too Late? Do the Risks of FemTech already Outweigh the Benefits?', European Symposium on Usable Security 2022, https://doi.org/10.1145/3549015.3554204; Alfawzan, N. and Christen, M. (2023), 'The future of FemTech ethics & privacy – a global perspective', *BMC Medical Ethics* 24, https://doi.org/10.1186/s12910-023-00976-z.

before sharing their personal health information with others'.[38] Similarly, in 2022, the FTC filed a lawsuit against data broker Kochava Inc, which was accused of commercializing sensitive data, including visits to domestic violence shelters and reproductive health clinics – data that could be used to determine a user's home address.[39] According to the FTC's data sample, time-stamped geolocation data from over 61 million unique mobile services was potentially affected. [40]

In 2022, 32 data brokers in the US were found to promote access to datasets containing data on millions of data subjects in the US labelled as 'actively pregnant', 'potentially pregnant' or 'shopping for maternity products', also offering large datasets on people using birth control products that have recently been banned by some US states.[41]

At the time of writing, there have been no examples of reproductive health-related information sold by data brokers being presented as evidence in court against those accused of having sought abortions in the US. However, US authorities have previously used digital evidence such as text messages and online search histories to enforce abortion laws.[42] In some US states, the evidence bar for abortion cases can be extremely low: geolocation data proving a visit to an abortion clinic alone could be considered strong enough evidence for a conviction.[43] Law enforcement is also increasingly using 'reverse warrants' focused on geolocation data or search histories (these are known as 'geofence warrants' or 'keyword warrants'). There are concerns that reverse warrants could be used to justify digital 'dragnets', identifying large numbers of potential abortion-seekers in the process.[44] There are also concerns that bounty hunters might use purchased geolocation data to track abortion patients and providers, motivated by new monetary rewards on offer in some US states.[45] Anti-abortion groups have also reportedly obtained data from brokers and used it to target women and abortion advocates with online disinformation, harassment and doxxing.[46]

[38] Federal Trade Commission (2021), 'FTC Finalizes Order with Flo Health, a Fertility-Tracking App that Shared Sensitive Health Data with Facebook, Google, and Others', press release, 22 June 2021, https://www.ftc.gov/news-events/news/press-releases/2021/06/ftc-finalizes-order-flo-health-fertility-tracking-app-shared-sensitive-health-data-facebook-google.

[39] Federal Trade Commission (2022), 'FTC Sues Kochava for Selling Data that Tracks People at Reproductive Health Clinics, Places of Worship, and Other Sensitive Locations', press release, 29 August 2022, https://www.ftc.gov/news-events/news/press-releases/2022/08/ftc-sues-kochava-selling-data-tracks-people-reproductive-health-clinics-places-worship-other. As of February 2024, this case was still pending. See Federal Trade Commission (2024), 'FTC v Kochava, Inc.', Cases and Proceedings, last updated 5 February 2024, https://www.ftc.gov/legal-library/browse/cases-proceedings/ftc-v-kochava-inc.

[40] Ibid.

[41] Wodinsky, S. and Barr, K. (2022), 'These Companies Know When You're Pregnant–And They're Not Keeping It Secret', Gizmodo, 30 July 2022, https://gizmodo.com/data-brokers-selling-pregnancy-roe-v-wade-abortion-1849148426; Ollove, M. (2022), 'Some States Already Are Targeting Birth Control', Stateline, 19 May 2022, https://stateline.org/2022/05/19/some-states-already-are-targeting-birth-control.

[42] Conti-Cook, C. (2020), 'Surveilling the Digital Abortion Diary', *University of Baltimore Law Review*, 50(10), pp. 1–76, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3666305.

[43] Wodinsky and Barr (2022), 'These Companies Know When You're Pregnant–And They're Not Keeping It Secret'.

[44] Riley, T. (2023), 'California lawmaker seeks to end to 'reverse warrants' that could pinpoint abortion seekers', Cyberscoop, 13 February 2023, https://cyberscoop.com/california-lawmaker-reverse-warrants-abortion.

[45] Lennard, N. (2021), 'The Supreme Court Will Not Save Us From the Decimation of Abortion Rights', The Intercept, 1 September 2021, https://theintercept.com/2021/09/01/texas-abortion-rights-sb8-supreme-court.

[46] Ascher-Schapiro, A. and Molone, A. (2023), 'U.S. abortion advocates face doxxing as data scavenged online', Context, 1 August 2023, https://www.context.news/digital-rights/us-abortion-advocates-face-doxxing-as-data-scavenged-online; Electronic Privacy Information Center (2024), 'Data Broker Helped Anti-Abortion Group Target Planned Parenthood Visitors, Wyden Letter Reveals', 13 February 2024, https://epic.org/data-broker-helped-anti-abortion-group-target-planned-parenthood-visitors-wyden-letter-reveals.

This evidence exemplifies the various ways through which reproductive health data sold by data brokers have been weaponized or hold the potential to be weaponized by public and private actors alike. In this case, a profit motive informed the design of highly data-extractive technologies (i.e. fem-tech apps) and the data generated by this design choice were exploited by both brokers and technology companies. Potential harms directly generated by such interactions include the prosecution of women seeking healthcare that is criminalized (e.g. information and services related to abortions, sexually transmitted infections, contraception, transgender healthcare, etc.), targeted disinformation about reproductive healthcare, and stalking, doxxing and even physical violence against patients, healthcare providers and activists.

In addition to the deliberate 'weaponization' of reproductive health data, there are also an array of indirect harms to consider, some rooted in mistrust and others in the consequences of increased dependence on online services, as availability of offline resources becomes increasingly limited. These indirect harms include wrongful convictions due to biased and unreliable data purchased from data brokers or retrieved from fem-tech applications, and discrimination by employers. Some data privacy advocates have raised concerns that, for example, an employer who opposes abortion could weaponize predictive employment algorithms to discriminate against candidates who have, or are suspected to have had, an abortion.[47] In such cases, the fear of being prosecuted may in itself restrict access to healthcare information and services, while, if certain data generated by fem-tech apps are used for research and to inform public policy, datasets risk encoding misrepresentation and bias.

In addition to the deliberate 'weaponization' of reproductive health data, there are also an array of indirect harms to consider, some rooted in mistrust and others in the consequences of increased dependence on online services.

The profit motive of private actors (in this case, both data brokers and technology companies) creates an enabling environment where the gendered security of those seeking reproductive healthcare is considered as a secondary or competing priority, as is the security of the data generated by fem-tech devices and apps itself. While technology companies and application developers depend on public opinion and users' trust, and thus have an incentive to improve their privacy policies when faced with public criticism, this is not the case for data brokers, for whom public warnings of the vast data they gather and the potential 'weaponizability' of this data might even serve as advertisement for the effectiveness of their services.

[47] Boodman, E., Tannow, T., Herman, B. and Ross, C. (2022), 'HIPAA won't protect you if prosecutors want your reproductive health records', STAT, 24 June 2022, https://www.statnews.com/2022/06/24/hipaa-wont-protect-you-if-prosecutors-want-your-reproductive-health-records.

4.3 'Weaponization' of LGBTIQ+ dating app data

Dating apps collect large amounts of personal information about their users and often share it with third parties, sometimes without adequately informing users about how these data are used and who can access them.[48] While app developers claim that these data are anonymized, users can be re-identified, not least because dating apps collect and can share geolocation data with third parties.[49] Similar to data about reproductive health, information about people's sexual orientation and dating behaviour has the potential for commercialization and 'weaponization', resulting in potential harms faced by individuals with different gender and sexual identities.

Both states that criminalize homosexuality and anti-LGBTIQ+ activists may use dating apps to spy on and persecute LGBTIQ+ individuals by exploiting apps' lack of privacy safeguards,[50] but also potentially by purchasing data collected and aggregated by data brokers who have commercial relations with app providers.

For example, the LGBTIQ+ dating app Grindr has faced criticism for its data privacy practices.[51]

The commercial availability of data collected by apps like Grindr is highly concerning due to its potential for 'weaponization' by state and non-state actors alike. For example, it could be used by third parties to target LGBTIQ+ individuals in religious communities or organizations where being outed could result in discrimination and other harms.

There are similar security concerns regarding the 'weaponization' of Grindr data by state actors. A Human Rights Watch report recently documented multiple cases of security agencies in the Middle East and North Africa (including those in Egypt, Iran, Iraq, Jordan, Lebanon and Tunisia) using dating apps such as Grindr to identify and prosecute LGBTIQ+ people.[52] In Egypt, for example, police have been reported to use features of Grindr to persecute gay men: Grindr's location feature, screenshots and messages, and even the mere presence of the app on an individual's phone can form part of 'debauchery' court cases.[53] In this context, location data retrieved from dating apps, or infiltration of the apps themselves (via fake profiles, for example) can enable harms like threats, assaults or even

[48] For example, see Heilweil, R. (2020), 'Tinder may not get you a date. It will get your data', Vox, 14 February 2020, https://www.vox.com/recode/2020/2/14/21137096/how-tinder-matches-work-algorithm-grindr-bumble-hinge-algorithms.
[49] Kaspersky Team (2021), 'Online dating and security', Kaspersky Daily, 16 July 2021, https://www.kaspersky.com/blog/mwc21-online-dating-apps/40628; Valentino de Vries, J., Singer, N., Keller, M. H. and Krolik, A. (2018), 'Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret', *New York Times*, 10 December 2018, https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html.
[50] A report published by Human Rights Watch in 2023 summarized digital targeting and its offline consequences for LGBTIQ+ people in the Middle East and North Africa. See Human Rights Watch (2023), *"All This Terror Because of a Photo"*.
[51] Franceschi-Bicchierai, L. (2021), 'Grindr Has Been Warned for Years About Its Privacy Issues', Vice Motherboard, 22 July 2021, https://www.vice.com/en/article/wx57nm/grindr-has-been-warned-for-years-about-its-privacy-issues; Singer, N. and Krolik, A. (2021), 'Grindr and OkCupid Spread Personal Details, Study Says', *New York Times*, Updated 14 October 2021, https://www.nytimes.com/2020/01/13/technology/grindr-apps-dating-data-tracking.html.
[52] Human Rights Watch (2023), *"All This Terror Because of a Photo"*.
[53] Brandom, R. (2018), 'Designing for the crackdown', The Verge, 25 April 2018, https://www.theverge.com/2018/4/25/17279270/lgbtq-dating-apps-egypt-illegal-human-rights; Human Rights Watch (2023), *"All This Terror Because of a Photo"*.

arrests of LGBTIQ+ people. The use of dating apps by to spy on and arrest LGBTIQ+ individuals is simultaneously an exploitation of an app's lack of privacy safeguards and an abuse of user expectations about how their data is used.

Following criticism from human rights organizations, Grindr introduced several design changes, including features like enabling users to unsend messages, block screenshots and send temporary images designed to expire. Grindr has also launched a 'Holistic Security Guide', which provides guidance on reducing the potential for harm across users' digital security, personal safety, and self-care and well-being.[54] The app also disabled the distance function in countries such as Egypt, Iraq, Nigeria, Russia and Saudi Arabia.[55] Most recently, following reports of LGBTIQ+ people being arrested in Egypt in early 2023, Grindr sent a warning to its users in the country, stating:

> We have been alerted that Egyptian police is actively making arrests of gay, bi, and trans people on digital platforms. (…) They are using fake accounts and have also taken over accounts from real community members who have already been arrested and had their phones taken.[56]

As dating apps like Grindr improve their data privacy practices in response to public scrutiny, and user awareness of data privacy risks associated with the use of such apps increases, police or government agents targeting LGBTIQ+ individuals may increasingly turn to data brokers to help identify individuals, circumventing data privacy practices by using geolocation data for de-anonymization. Even if dating apps address gaps in their practices and data brokers refrain from collecting sensitive data in the future, these measures will come too late for users or former users whose data have already been collected and could still be up for sale.

[54] King, J. P. (2019), 'Grindr rolls out new features for countries where LGBTQ identity puts users at risk', Washington Blade, 13 December 2019, https://www.washingtonblade.com/2019/12/13/grindr-rolls-out-new-features-for-countries-where-lgbtq-identity-puts-users-at-risk. Grindr's approach to holistic security is led by Grindr For Equality, which 'has partnered with LGBTIQ+ and health activists and organizations around the world—as well as our users—on a Holistic Security Guide that encompasses multiple areas of user safety'. See Grindr For Equality (undated), *Grindr Holistic Security Guide*, https://assets.website-files.com/641dc6058ca7b72a1422b5d7/6436c2bf48bde3005f39e0eb_G4E-HolisticSecurityGuide-English.pdf (accessed 29 May 2024).
[55] @Seppevdpll (2018), 'It is Still Possible to Obtain the Exact Location of Millions of Men on Grindr', Queer Europe blog, 13 September 2018, https://www.queereurope.com/it-is-still-possible-to-obtain-the-exact-location-of-cruising-men-on-grindr.
[56] Al Jazeera (2023), 'LGBTQ+ dating app Grindr warns Egypt users of police-run accounts', 25 March 2023, https://www.aljazeera.com/news/2023/3/25/lgbtq-dating-app-grindr-warns-egypt-users-of-police-run-accounts.

05 Private sector–state interaction in response to gendered cyber harms

The choice and uses of technologies for responding to cybercrime are significant, as is the choice of supplier or provider.

The global rise in cybercrime and cybercriminal threat actors – and the subsequent expansion of cybercrime victims to include individuals, communities, companies and sometimes even entire governments[57] – has been coupled with an increasing global awareness of the gendered dimensions of the cybercrime life cycle,[58] from gender-based victimization to gender-disaggregated data on the impact of cybercrime. At the same time, there are important global efforts towards ensuring

[57] INTERPOL (2023), 'Urgent global response needed for "insidious" cybercrime', 16 October 2023, https://www.interpol.int/en/News-and-Events/News/2023/Urgent-global-response-needed-for-insidious-cybercrime.
[58] For example, see INTERPOL (2023), *Policing with a gender perspective*, compendium, https://www.interpol.int/News-and-Events/News/2023/Policing-with-a-gender-perspective; Millar, K., Shires, J. and Tropina, T. (2021), *Gender approaches to cybersecurity: design, defence and response*, United Nations Institute for Disarmament Research, https://unidir.org/wp-content/uploads/2023/05/Gender-Approaches-to-Cybersecurity_Digital_Final.pdf.

cybercrime responses are gender-sensitive and some promising developments.[59] Globally, law enforcement both redesigns and readapts technologies developed in the private sector for this purpose, such as online reporting platforms and tools for monitoring harmful content.

However, significant gaps persist in responses to cybercrime.[60] In some cases, state and police responses to gendered cyber harms can mitigate harm. In others, they can lead to secondary harms. The choice and use of technology for responding to cybercrime are significant,[61] as is the choice of supplier or provider.

This chapter considers case studies that interact with, but are not the direct result of, profit motives in technology design: first, responses to digital sex crimes in South Korea; and second, responses to gender-based violence in India. The chapter explores the redesign and readaptation of technologies to enable state and police responses to gendered cyber harms, focusing on the potential implications of design and deployment choices.

5.1 Investigating digital sex crimes and spycam abuse in South Korea

The proliferation of digital sex crimes in South Korea presents a case study for exploring the ways in which digital technologies that are designed to alleviate or prevent gendered cyber harms and enhance cybersecurity can mitigate and/or exacerbate such harms in new, predictable or unpredictable ways. The technologies in question are used for monitoring harmful and/or illegal content.

In recent years, key markers of internet connectivity in South Korea (i.e. mobile connections, social media users and internet connection speeds) have increased. So too have technology-facilitated abuses, most notably '*molka*' (몰카) crimes – meaning digital sex crimes involving the use of hidden spycams to capture intimate or private images without knowledge or consent, and the dissemination of this non-consensually captured content via public or private channels.[62] In the 'global epicentre'[63] of spycam abuses, prosecutions for sex crimes involving illegal filming in South Korea rose 11-fold between 2008 and 2017,[64] but survivors (for the most part, women and girls)[65] reportedly face multiple barriers to reporting,

[59] For example, in the UK: UK Ministry of Justice and The Rt Hon Lucy Frazer KC MP (2019), "Upskirting' now a specific crime as bill receives Royal Assent', press release, 12 February 2019, https://www.gov.uk/government/news/upskirting-now-a-specific-crime-after-bill-receives-royal-assent; UK Ministry of Justice and Laura Farris MP (2024), 'Government cracks down on 'deepfakes' creation', press release, 16 April 2024, https://www.gov.uk/government/news/government-cracks-down-on-deepfakes-creation.
[60] In many cases, while an anti-cybercrime stakeholder (for example, a prosecutor or investigator) may be aware of gender considerations in implementing anti-cybercrime responses, responses themselves can be an additional source of harm. For further reading, see Millar, Shires and Tropina (2021), *Gender approaches to cybersecurity*.
[61] The definition of cybercrime is also highly politicized. As Chapter 4 explores, police forces have used dating apps to target LGBTIQ+ individuals on morality charges. There is a global trend of state authorities abusing cybercrime laws to endanger human rights and fundamental freedoms.
[62] Barr, H. (2021), *My Life Is Not Your Porn: Digital Sex Crimes in South Korea*, report, New York: Human Rights Watch, https://www.hrw.org/report/2021/06/16/my-life-not-your-porn/digital-sex-crimes-south-korea.
[63] Yi, B. L. (2020), 'Untouched yet ruined: toll of S.Korea spycam epidemic', Thompson Reuters Foundation, 13 January 2020, https://news.trust.org/item/20200113002724-jflg1.
[64] Data from the Korean Institute of Criminology cited in Barr (2021), *My Life Is Not Your Porn*.
[65] See Barr (2021), *My Life Is Not Your Porn*. Barr writes that 'the overwhelming majority of the people targeted in digital sex crimes are women – 80 per cent in spycam cases. The overwhelming majority of perpetrators are male; in 2016, 98 per cent of perpetrators in spycam cases were men.'

justice and recovery. These range from social barriers – including social stigma and reputational harm – to institutional ones created and exacerbated by political leaders,[66] police, prosecutors and legislators.[67] These barriers are a significant cybersecurity risk, as expectations of privacy are not met, and both technical and social vulnerabilities are exploited by malicious actors. They are a gendered risk because of the disproportionate impact of *molka* crimes on women and girls.[68]

South Korean legislators have enacted several measures[69] to address, investigate and prevent spycam abuse, although activists have criticized these measures as insufficient,[70] unsustainable[71] or as enablers of surveillance without safeguards.[72] Technology companies like Google have also been criticized for exacerbating harm through their 'inadequate' reporting system.[73] Gaps in these measures, and in enforcement, justice and survivor support, are reportedly often filled by private sector companies, private individuals and civil society groups.[74] This chapter outlines two (related) state-led initiatives for monitoring digital sex crimes.

The Digital Sex Crimes monitoring unit/taskforce at the Korea Communications Standards Commission (KCSC) was founded in 2019 with a mandate to monitor domestic and foreign websites for Korean-language hashtags that may reference images and videos captured without consent.[75] With KCSC's regulatory power behind it, the taskforce can force South Korean sites to take down images and videos. But for content hosted on overseas servers, it can only request that foreign operators remove it.[76] In 2022, the taskforce was recommended by the Ministry of Justice to take further steps to delete and block illegal videos, such as preventing abusive videos

[66] Current South Korean president Yoon Suk-yeol's electoral platform was widely criticised as anti-feminist. One of his campaign promises was to clamp down on 'false' accusations of sex crimes. Since taking office, the number of 'false' accusation investigations has increased. See Jung, H. (2023), 'Are South Korea's New Policies Silencing Rape Survivors?' Foreign Policy, 30 June 2023. https://foreignpolicy.com/2023/06/30/south-korea-rape-survivors-feminism-false-accusations.

[67] The high-profile case of K-pop star Jung Joon-Yong's crimes against 12 women is an example of multiple barriers to bringing charges against known offenders. See Bicker, L. (2021),"I was humiliated': The continuing trauma of South Korea's spy cam victims', BBC News, 16 June 2021, https://www.bbc.co.uk/news/world-asia-57493020.

[68] See Barr (2021), *My Life Is Not Your Porn*.

[69] See the agreed actions of a pan-government council on strengthening protections against digital sex crimes, convened in 2020: Republic of Korea Ministry of Gender Equality and Family (2020), 'Harsher Punishment and Stronger Protection against Digital Sex Crimes', 23 April 2020, https://www.mogef.go.kr/eng/pr/eng_pr_s101d.do?mid=eng001&bbtSn=707003. For further information on how South Korea filters 'illegal' content hosted on overseas-based websites, see Morgus, R., Sherman, J. and Nam, S. (2019), 'Analysis: South Korea's New Tool for Filtering Illegal Internet Content', New America blog, 15 March 2019, https://www.newamerica.org/cybersecurity-initiative/c2b/c2b-log/analysis-south-koreas-sni-monitoring.

[70] Nguyen, E. (2021), 'From Schools to Motels, Nowhere is Safe from Spy Cams in South Korea', Human Rights Watch, 4 November 2021, https://www.hrw.org/news/2021/11/04/schools-motels-nowhere-safe-spy-cams-south-korea.

[71] Amnesty International (2023), *Republic of Korea: Submission to the UN Committee on the Elimination of Discrimination Against Women*, https://www.amnesty.org/en/wp-content/uploads/2023/01/ASA2563912023ENGLISH.pdf.

[72] Rosen, K. R. (2022), 'In South Korea, women are fighting to end digital sex crimes', Coda Story, 20 December 2022, https://www.codastory.com/authoritarian-tech/molka-digital-sex-crimes-south-korea.

[73] Amnesty International (2022), 'South Korea: Online sexual abuse content proliferates as survivors blame Google failings', 8 December 2022, https://www.amnesty.org/en/latest/news/2022/12/south-korea-online-sexual-abuse-content-proliferates-as-survivors-blame-google-failings.

[74] These organizations range from informal, student-run groups like Project ReSet (Reporting Sexual Exploitation on Telegram), whose work led to the 'Nth Room' arrests, to private companies offering 'digital laundry' services or the removal of tracking tools. See Rosen (2022), 'In South Korea, women are fighting to end digital sex crimes'; France 24/AFP (2017), 'Battle of the online sex crimes in high-tech S. Korea', 13 October 2017, https://www.france24.com/en/20171013-battle-online-sex-crimes-high-tech-korea; Kim, J-M. (2020), "Digital undertakers' are busier than ever', Korea JoongAng Daily, 5 April 2020, https://koreajoongangdaily.joins.com/2020/04/05/industry/Digital-undertakers-are-busier-than-ever/3075699.html.

[75] France 24/AFP (2019), 'S. Korean regulators in constant search for porn', 20 November 2019, https://www.france24.com/en/20191120-s-korean-regulators-in-constant-search-for-porn.

[76] Ibid.

from dissemination by blocking access.[77] The taskforce head also noted that survivors directly contact them to deal with cases, although the KCSC reportedly also operates a separate victim support centre.[78] However, no information is available publicly about the specific monitoring technologies used by the taskforce.[79]

While the KCSC's taskforce has a nationwide mandate, local police agencies and local governments have also pioneered their own monitoring initiatives. For example, the Seoul Metropolitan Government's (SMG) AI-based monitoring system[80] was announced in early 2023 to replace manual monitoring, based at the city's digital sex crime centre. AI-enabled monitoring in this programme reportedly identifies sexually exploitative material.[81] The programme has a mandate for automatically deleting content and 'blocking circulation' at the source.[82] Official reporting on the legislative mechanism underlying the programme's mandate for deleting content appear vague, as is information on the development of the AI program itself, which is reported to have been developed by the Seoul Institute for Technology.[83]

Both initiatives discussed were designed to address a specific gendered cyber harm through the use of monitoring technologies by specialized units and programmes. However, there is a danger of the solution leading to secondary harms for survivors.

The KCSC's lack of transparency has been criticized by civil society organizations,[84] particularly the body's enforcement of 'vaguely defined standards and broad discretionary power' in the sub-commission on internet communications – that enables commissioners to 'make politically, socially and culturally biased judgements', which may lack a legislative rationale.[85] Furthermore, there appears to be little publicly available information about the selection, development and use of monitoring technologies by either the KCSC's taskforce and the SMG's unit. This lack of public information may not necessarily suggest an absence of independent expert oversight in the technology's selection, design, implementation and auditing.[86]

The definition of 'digital sex crime' materials is enshrined in South Korean legislation,[87] but the interpretation and operationalization of these definitions by state, police, suppliers and implementers may not be fully aligned, especially

[77] Ministry of Justice of The Republic of Korea (2022), 'Digital Sex Crimes Task Force Team – Expert Committee Activities and Achievements', 25 April 2022, p. 20. On the 2020 government announcement on the 'Eradication of Digital Sex Crimes' see also Korea Communications Commission (2020), 'Annual Report', p. 98.
[78] France 24/AFP (2019), 'S. Korean regulators in constant search for porn'.
[79] At the time of writing, and based on the author's research, which is limited by language.
[80] This initiative appears to be part of a new wave of initiatives trialling the use of AI and facial recognition technology to improve monitoring for sexually exploitative content. See Yonhap (2021), 'Police to consider disclosing identities of online sexual abuse material buyers', *Korea Herald*, 13 December 2021. https://www.koreaherald.com/view.php?ud=20211213000570.
[81] Seoul Metropolitan Government (2023), 'Seoul cracks down on digital sex crimes using AI technology', press release, 30 March 2023, https://english.seoul.go.kr/seoul-cracks-down-on-digital-sex-crimes-using-ai-technology.
[82] Ibid.
[83] Secondary reporting on this subject is also limited. For instance, see Boram, P. (2023), 'Seoul city uses AI technology to monitor around clock', Yonhap News Agency, 29 March 2023, https://en.yna.co.kr/view/AEN20230329005200315.
[84] York, J. C. and Reitman, R. (2011), 'In South Korea, the Only Thing Worse Than Online Censorship is Secret Online Censorship', Electronic Frontier Foundation Deeplinks blog, 6 September 2011, https://www.eff.org/deeplinks/2011/08/south-korea-only-thing-worse-online-censorship; Freedom House (2022), 'Freedom on the Net 2022: South Korea', https://freedomhouse.org/country/south-korea/freedom-net/2022.
[85] York and Reitman (2011), 'In South Korea' (as cited in Freedom House (2022), 'Freedom on the Net 2022: South Korea').
[86] There may be important reasons that KCSC, SMG and the Seoul Institute for Technology do not publicly disclose how their monitoring technologies work, such as cybersecurity considerations.
[87] See Republic of Korea Ministry of Gender Equality and Family (2020), 'Harsher Punishment and Stronger Protection against Digital Sex Crimes'.

those regarding exactly what constitutes illegal, harmful or exploitative material, and the minimum benchmark for automatic deletion or platform notification. The implications of this misalignment could range from the wrongful criminalization of legitimate content to biased decision-making and outcomes.

There is the added challenge of ensuring that mechanisms for updating and reviewing identifiers consider developments in technology (particularly AI-generated deepfakes). The use of digital sex crime images and videos to train AI models more generally raises urgent questions about whether sufficient safeguards (such as data input controls and auditing measures) are implemented, particularly as the design and implementation of technologies relying on harmful datasets can lead to the reflection and exacerbation of existing systemic biases.[88] If this was the case, a potential secondary harm faced by survivors would be inaccurate notification, which could result in (re)traumatization and psychological distress, regardless of the outcome.

Technologies designed and readapted to deliver gender-transformative cybersecurity must be responsive to potential harms (and harm mitigation measures) across the life cycle of design, deployment and review.

Monitoring is just one point in the life cycle of measures addressing digital sex crimes. Gender-sensitive monitoring is not in itself a safeguard against compounded harms at other points in the cycle. Human Rights Watch reports systemic failings by police (such as victim-blaming or jokes about intimate images) and judges (with trends towards lenient sentencing, and preferences for issuing fines rather than imprisonment) compounding survivors' existing difficulties in seeking criminal justice.[89] The impact of monitoring interventions could be undermined by weaknesses elsewhere in the country's political and justice system, leading to inadequate outcomes at best and compounded harms at worst.

In this case, technologies designed and readapted to deliver gender-transformative cybersecurity must be responsive to potential harms (and harm mitigation measures) across the life cycle of design, deployment and review. This approach is consistent with a sociotechnical approach to cybersecurity more generally, which considers both the security of technologies and the security needs of users. Partnerships between private sector actors, law enforcement and academia have the potential to deliver results, but also to compound harms.

[88] This gendered harm is well documented in the literature. See, for example, Schwartz, R. et al. (2022), *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, National Institute of Standards and Technology, Special Publication 1270, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf.
[89] Ngyuen, E. and Barr, H. (2020), 'Thinking Beyond Punishment to Combat Digital Sex Crimes in South Korea', Human Rights Watch, 20 May 2020, https://www.hrw.org/news/2020/05/20/thinking-beyond-punishment-combat-digital-sex-crimes-south-korea. See also Mackenzie, J. (2022), 'As South Korea abolishes its gender ministry, women fight back', BBC News, 14 December 2022, https://www.bbc.co.uk/news/world-asia-63905490.

5.2 Reporting gender-based violence in India

South Korea is not alone in trialling technological solutions to gender-based violence. India is another country that has implemented several initiatives to improve online reporting of such crimes. Data from the Indian National Crime Records Bureau (NCRB) reports a 'consistent year-on-year rise' in gender-based violence from 2016 to 2021, with the exception of 2020.[90] Meanwhile, the UN Sustainable Development Group referred to gender-based violence in India during COVID-19 as a 'shadow pandemic'.[91]

While much gender-based violence is experienced by survivors offline, there are rising concerns about technology enabled violence, which can both transform offline harms and lead to new types of cyber harm. A report by the University of Chicago and the International Center for Research on Women suggests that 'male dominance in online spaces and gendered cultural norms often make the internet inhospitable for women' in India. [92] For more than a decade, activists and academics have mapped gendered cyber harms experienced by marginalized or minoritized gender identities in online spaces in India,[93] as well as the measures used to mitigate and prevent online abuse (as explored below).[94] These studies map out a landscape with deeply entrenched barriers to reporting gender-based violence, whether experienced online, offline or both.

The Indian government, states and local police agencies have pledged to address gender-based violence through reporting mechanisms. The national Scheme for Cyber Crime Prevention against Women and Children (CCPWC), pioneered by the Ministry of Home Affairs and the Ministry of Women and Child Development, aims to 'have an effective mechanism to handle cybercrimes against women and children in the country'. [95] One of the scheme's main features was the 2018 launch of an online cybercrime-reporting platform for complaints relating to child sexual abuse material or sexually explicit content.[96] The National Commission for Women (NCW) supports this initiative and others,[97] including a 24/7 helpline 'to help women facing domestic violence', and a WhatsApp helpline launched in April

[90] Pandey, G. (2022), 'Rising crimes against Indian women in five charts', BBC News, 13 September 2022, https://www.bbc.co.uk/news/world-asia-india-62830634.
[91] United Nations Sustainable Development Group (2020), 'Shadow Pandemic: UN India responds to uptick in violence against women and girls during COVID-19', 9 December 2020, https://unsdg.un.org/latest/stories/shadow-pandemic-un-india-responds-uptick-violence-against-women-and-girls-during.
[92] NORC at the University of Chicago and the International Center for Research on Women (2022), *Technology-Facilitated Gender-Based Violence In India*, Washington, DC: United States Agency for International Development, https://www.icrw.org/wp-content/uploads/2021/09/USAID-TFGBV-India.pdf.
[93] Kovacs, A., Padte, R. K. and S. V. Shobha (2013), *'Don't Let It Stand!' An Exploratory Study of Women and Verbal Online Abuse in India*, New Delhi: Internet Democracy Project (April 2013), https://cdn.internetdemocracy.in/idp/assets/downloads/reports/women-and-verbal-online-abuse-in-india/Internet-Democracy-Project-Women-and-Online-Abuse.pdf.
[94] Chandrashekar, R. (2017), *Policing online abuse or policing women? Our submission to the United Nations on online violence against women*, report, New Delhi: Internet Democracy Project, 7 November 2017, https://internetdemocracy.in/policy/un-srvaw-report.
[95] Press Information Bureau, Government of India Ministry of Home Affairs (2019), 'Cyber Crime prevention against Women and Children', press release, 8 January 2019, https://pib.gov.in/Pressreleaseshare.aspx?PRID=1559115.
[96] Ibid.
[97] Government of India, National Commission for Women (undated), 'Cyber Crime Prevention Against Women and Children (CCPWC)', http://ncw.nic.in/ncw-cells/legal-cell/new-bills-laws-proposed/cyber-crime-prevention-against-women-and-children-ccpwc.

2020.[98] In the last 10 years, technology companies have also launched personal safety apps for smartphone users, with features for location-sharing, sending emergency alerts and notifications of unsafe locations.[99] This case study explores the use of WhatsApp chats for reporting gender-based violence to police.

As journalist Mahima Jain reports, the 'efficacy' of personal safety apps is hindered by various factors, including regional differences in language, technical glitches and complicated registration processes. As a result, 'the apps remain niche… none have been widely adopted'.[100] Jain interviewed the director of the women's safety wing in the Telangana State Police, who explained that 'WhatsApp has emerged as the most-used platform for women to seek help or make complaints of harassment'.[101] WhatsApp reportedly works closely with law enforcement agencies in India and notes that:

> In coordination with the Government of India we have provided law enforcement training on how to use WhatsApp as a resource in their community, respond to citizens on WhatsApp and make legal requests to WhatsApp in the process of investigating a crime.[102]

In 2022, the company launched a dedicated safety hub for users in India,[103] which includes resources on preventing abuse and supporting cybersecurity.

WhatsApp is used by local police in India for reporting in different capacities. For example, dedicated teams in the Telangana State Police handle complaints received via WhatsApp by contacting the complainant and collecting evidence. Over 40 per cent of reports picked up by the these teams are collected via WhatsApp messages.[104] Complaints include 'non-heinous' acts committed and experienced both offline (such as public harassment) and online (such as 'lewd comments made on social media').[105] In 2017, local police in Pune initiated a WhatsApp group called 'BuddyCop' to provide' immediate access to police' in the event of gender-based violence. Within six months of the group being launched, 100,000 women had registered and around 750 groups formed.[106] Pune also launched a WhatsApp helpline number (not a specialized chat, as in Telangana state) in July 2023, 'especially for women's safety and security', that forwards messages to the relevant police stations.[107]

[98] Gupta, S. (2021), 'Creating Safe Spaces: NCW Launches 24/7 Helpline For Women Facing Violence', The Logical Indian, 27 July 2021, https://thelogicalindian.com/trending/ncw-launches-24-hour-helpline-number-for-women-29898.
[99] The use of geolocation features for personal safety apps adds further complexity to the cases explored in Chapter 4 and serves to emphasize that technology features can be used or weaponized in different contexts, by different actors, for different outcomes. See Singh, K. (2023), ''Popular' women's safety apps in India', *Times of India*, 11 April 2023, https://timesofindia.indiatimes.com/gadgets-news/popular-womens-safety-apps-in-india/photostory/99388808.cms.
[100] Jain (2021), 'How WhatsApp became a tool for Indian police to fight harassment'.
[101] Ibid.
[102] WhatsApp Help Center (undated), 'Ensuring User Safety in India on WhatsApp', https://faq.whatsapp.com/650453616475879/?locale=en_US.
[103] Ibid; Balakumar K. (2022), 'WhatsApp launches 'Safety in India' hub - A resource for online safety', Tech Radar, 22 February 2022, https://www.techradar.com/news/whatsapp-launches-safety-in-india-hub-a-resource-for-online-safety.
[104] Jain (2021), 'How WhatsApp became a tool for Indian police to fight harassment'.
[105] Ibid.
[106] Nowrojee, S. and Shebi, K. (2019), *Working Together for Girls' and Women's Safety in Public Spaces*, 3D Program for Girls and Women, https://the3dprogram.org/content/uploads/2019/06/3D-Program-Public-Safety-report-June-2019.pdf, p. 37.
[107] Desphpande, S. (2023), '274 complaints received on WhatsApp number in five days; 192 actionable: Pune police', *Hindustan Times*, 16 July 2023, https://www.hindustantimes.com/cities/pune-news/pune-police-receives-274-complaints-within-5-days-of-launching-women-s-safety-whatsapp-number-101689523853571.html.

Political and police interest in innovating reporting mechanisms is a positive development, but technology enabled responses to gendered harms (whether experienced in cyberspace or offline) are by no means a fix-all solution. Technology platforms may not necessarily alleviate existing barriers to reporting gender-based violence, despite the relatively widespread access to and use of smartphones and data. Existing barriers include a lack of awareness about how and where to report,[108] as well as social stigmas around reporting violence, harassment and abuse. Additionally, as 2018 data suggests, 71 per cent of respondents (adolescent girls from low-income households) 'do not feel confident to approach the police in a case of harm'.[109] Gendered gaps in digital literacy and access to devices – not to mention a lack of data collected on non-binary or genderqueer individuals – further complicate this picture.[110] Only 29 per cent of women in India actually own a smartphone, compared to 48 per cent of men.[111] Moreover, the percentage of women who own a device to which they have sole and exclusive access and use (i.e. a device they do not share with family members) is unclear. This lack of access poses a serious barrier to the efficacy and coverage of well-intended initiatives such as WhatsApp helplines, and underlines the importance of adopting an intersectional perspective when mapping these barriers.

Technology platforms may not necessarily alleviate existing barriers to reporting gender-based violence, despite the relatively widespread access to and use of smartphones and data.

Against this backdrop, using WhatsApp to facilitate reporting gender-based violence raises questions around how the *choice* of technology platform affects gendered cyber harms. First, while WhatsApp may be a suitable fit for police use – with end-to-end encryption and a variety of features tailored to threats faced specifically by users in India[112] – there is an apparent lack of publicly available information[113] on how identifiable chat data is stored and safeguarded by key stakeholders (i.e. local police agencies and organizations such as the NCW receiving complaints via WhatsApp).[114] In addition, information on how sensitive data, images or videos (for instance, abusive messages shared by a complainant with a local police force using a WhatsApp chat helpline) would be stored or protected is either not apparent or unavailable.[115]

[108] NORC at the University of Chicago and the International Center for Research on Women (2022), 'Case Study: Technology-facilitated Gender Based Violence in India'.
[109] Gupta, S. (2018), '10 Things Adolescent Girls Had To Say About Gender-Based Violence And Their Experience of Reporting Them', Feminism in India blog, 25 December 2018, https://feminisminindia.com/2018/12/25/adolescent-girls-gender-based-violence-experience.
[110] This is because most data appear to be disaggregated into two categories: men and women.
[111] Jeffrie, N. (2023), *The Mobile Gender Gap Report 2023*, London: GSMA, https://www.gsma.com/r/wp-content/uploads/2023/07/The-Mobile-Gender-Gap-Report-2023.pdf, p. 44.
[112] WhatsApp Help Center (undated), 'Ensuring User Safety in India on WhatsApp'.
[113] This may not necessarily be the policy or process itself (for privacy and security reasons, details may not be available to the public) but instead a statement of principles, safeguarding, etc.
[114] At the time of writing. Further research and interviews with key stakeholders could help in mapping the full reporting life cycle using WhatsApp chats or helplines.
[115] At the time of writing.

A second area of concern lies in the difference between automated chat functions (available for users of 'WhatsApp for Business') and personal chat functions. The former generally rely on a series of key-term identifiers to automate responses and determine escalation. For instance, in reference to the Telangana case, it is unclear what the key-term identifiers are that determine a minimum benchmark for action. While an offence may be formally defined in legislation, decisions to act on suspected offences are influenced by a patchwork of personal, social, cultural and political factors, and may differ from state to state, and person to person.

Reporting is just one of many measures for mitigating gendered harms. WhatsApp's effectiveness as a reporting mechanism is contingent on cultural, institutional and legislative realities, which can enable better responses but also open the door to potential abuse. Accurate data on reporting (and barriers to reporting), disaggregated nationally and by gender, are key. Further research could include interviews with law enforcement stakeholders and aim to map such barriers to better understand how platforms like WhatsApp can be used to overcome them.

In both of the case studies presented in this chapter, despite innovations and good intentions, the efficacy of the solutions presented is uncertain and there is a clear potential for secondary or unaddressed harms. Nonetheless, both studies demonstrate actions that can be taken for countering and mitigating gendered cyber harms – an important step towards gender-transformative cybersecurity.

06 Conclusion and recommendations

Profit motives and other incentives could be harnessed to encourage private sector actors to redesign their technologies to counter and mitigate gendered cyber harms.

Cyber insecurity can be profitable. If companies design technologies without gender-sensitive (cyber)security measures, third parties (including law enforcement, private actors and others) can exploit design choices to serve their own incentives and motivations. But profit motives and other incentives can also be harnessed to encourage private actors to redesign their technologies to counter and mitigate harms. While some technologies may not be gender-sensitive in their design, they can be used to enhance gender-transformative cybersecurity (for instance, by being implemented to directly prevent, monitor and respond to gendered harms).[116] This phenomenon is particularly significant when technologies are readapted and deployed by state and police actors, as explored in Chapter 5.

Each case study provides a commentary on individual experiences of cybersecurity and the gendered cyber harms that individuals may be confronted with. Chapter 3 documents cyber harms faced by queer social media users in Nigeria and South Africa, while the case studies in Chapter 4 exposes how these harms can be generated through the commercialization and 'weaponization' of sensitive data. Chapter 5 shows how technologies can be used to bolster cybersecurity and combat gender-based violence.

Gendered cyber harms and cyber insecurity, while experienced individually, pose a global security challenge. In cyberspace, insecurity at an individual or community level can bear global, exponential repercussions. Technologies developed and

[116] However, this does not mean that technologies are not themselves gendered.

deployed by private actors have a global impact, and technology companies are increasingly involved in developing cybersecurity solutions to advance a stable, peaceful cyberspace in which all people can participate safely and securely.

To this end, the following general recommendations for countering and mitigating harms, and incentivizing gender-transformative cybersecurity, are aimed at private sector stakeholders developing or deploying technologies (including technology companies, but also data brokers). Other stakeholders can and should seek to assist in their implementation.

Recommendations

Critically evaluate data sharing and cooperation with state entities, adopting a human rights and gender perspective to map potential harms.
In exploring data sharing agreements with state entities (such as law enforcement and national security agencies), private actors need to consider the following questions:

- Is cooperation necessary according to local laws?
- What gendered risks and harms might such cooperation lead to?
- How can these risks be managed or limited?
- What compromises are being made between market incentives (e.g. expanding the reach of a product or identifying new customers) and data privacy, especially for data relating to gender and sexual orientation?

These considerations should be rooted in international standards and principles, such as the UN Guiding Principles on Business and Human Rights. For example, data brokers could pledge to not have a commercial relationship with organizations or industries that hold discriminatory, exclusionary and harmful stances on reproductive healthcare.

Assess the efficacy of user privacy and data sharing settings, and the accessibility and ease of changing those options, adopting a gender perspective to map potential harms.
Private actors should seek to identify where data collection and sharing are necessary for the functioning of a product, and assess the actual added value of those functions (e.g. for advertising) alongside potential risks – particularly in instances where additional data collection brings extra revenues.

When collecting or handling sensitive data, private actors should implement best practices in data minimization (such as storage time and location, anonymization, etc.); allow pseudonymous or anonymous access to services; stop behavioural tracking; install end-to-end encryption by default; and refrain from collecting any location-based information. For example, healthcare app providers could refrain from collecting geolocation data tracking visits to abortion clinics and other similarly stigmatized healthcare facilities. Compliance with regulatory standards on data privacy is an important lever in this case. For private companies, compliance with privacy regulation is necessary and can mitigate potential risks. Regulatory standards on data privacy should also be applied when private companies are considering data sharing agreements with state entities such as law enforcement and national security agencies.

Map technological relationships with commercial partners and potential risks.
Nearly all actors in the private sector depend on, and operate in, a complex and often poorly understood web of technology and data. Such actors should dedicate resources to mapping and evaluating their technological relationships, including direct and indirect data flows, from a gender-sensitive perspective, with the aim of identifying and mitigating gendered risks. For example, mapping data dependencies (both direct and indirect) between social media platforms is essential for mitigating potential harms faced by certain social media platform users, who may have concerns about their content and personal information moving between platforms without their knowledge or consent.[117]

Implement additional technical features and mitigations that enable users to reduce risks.
Technology designers should proactively identify high-risk situations and contexts faced by their users, and design and implement features that enable users to reduce risks. Such features may include allowing the user to disable location services. Additional features and mitigations should be incorporated into design scoping, and assessed as part of the overall profitability of the product – ideally as part of a 'security-by-design' process. The commercial value of additional security features should also be considered, particularly as the product may attract more users if there are strong security features. Finally, information about opting-out and reducing risk must be proactively, accessibly and regularly shared with users (for example, via a clearly labelled reporting function).

Incorporate user experiences and feedback into technology design, redesign and readaptation.
'Building in' gender sensitivity at all points of the technology design process is essential for mitigating gendered harms – especially when the technologies themselves may be deployed by third parties to mitigate harms. Particular attention should be given to how the same technologies are accessed and used by people of different genders, and to the efficacy of cybersecurity measures for affected individuals. This can be partially achieved through incorporating user experience and feedback, whether through research, consultations or demonstrations.

Without sufficient safeguards and gender-sensitivity from the outset – drawing directly from individual experiences and implementing context – technologies can intentionally or inadvertently further entrench harms faced by marginalized groups (for example, through the over-censorship of online content or over-surveillance).

[117] This was investigated in the case studies in Chapter 3 and is based on testimony from an interviewee.

Build internal, independent gender expertise and connect to international networks for best practices.

Technology companies should nominate internal gender champions with sufficient seniority and independence from commercial decision-making processes. These champions should be supported with sustainable resources and training, and be authorized to develop information sharing networks internally. The latter is an important step towards incorporating gender into broader corporate social responsibility commitments. By connecting to existing international, multi-stakeholder networks, events and initiatives,[118] private sector actors can share and further develop best practices, and facilitate public-private partnerships for the redesign and readaptation of technologies for the purpose of harm mitigation.

[118] These might include attending recurring high-level conferences on gender equality; contributing to taskforces and consultations (for example, those run by civil society organizations and research institutes); and joining international coalitions for action (such as multi-stakeholder pledges to address gendered cyber harms).

About the authors

Isabella Wilkinson is a research fellow in the Digital Society Initiative at Chatham House. Isabella's work covers security and governance issues relating to cyberspace and technology, including international cyber and technology governance and processes; the (online) information environment; and advancing responsibility and gender inclusivity.

Julia-Silvana Hofstetter is a senior adviser at the ICT4Peace Foundation and a non-resident research fellow in the AI Security Initiative at UC Berkeley's Center for Long-Term Cybersecurity. Her areas of expertise include digital peacebuilding; disinformation and hate speech in armed conflict; humanitarian data governance; and human-centric cybersecurity.

James Shires is the co-director of both the European Cyber Conflict Research Incubator (ECCRI CIC) and the European Cyber Conflict Research Initiative. With ECCRI CIC, James leads the Google.org European Cybersecurity Seminars programme, a multi-year initiative to expand AI and cybersecurity education across Europe.

Mardiya Siba Yahaya is digital sociologist, researcher and community movement builder, whose work investigates the implications of technology-facilitated surveillance and datafied societies on minoritized communities in the Global South.

Acknowledgments

This research paper is the second in a series produced as part of a Chatham House project on gender and cybersecurity funded by the UK Foreign, Commonwealth and Development Office (FCDO).

The authors are grateful for the support of the FCDO Gender, Peace and Security team within the Integrated Security FUND (previously Conflict, Security and Stability Fund) in enabling this project.

We also offer sincere thanks to the project's advisory group and the members of the International Security Programme at Chatham House for their feedback and support, particularly Amrit Swali, Georgia Cole and Joyce Hakmeh.

We would also like to thank Chris Matthews in the Communications and Publishing department at Chatham House for editing the paper, and are extremely grateful to the anonymous peer reviewers for their feedback.

Chatham House does not express opinions of its own. The opinions expressed in this publication are the responsibility of the author(s).

Cover image: People gather for a concert during the Soweto Pride festival, 25 September 2021,
Soweto, South Africa.
Photo credit: Copyright © Luca Sola/AFP via Getty Images

ISBN 978 1 78413 609 3
DOI 10.55317/9781784136093

Cite this paper: Wilkinson, I., Hofstetter, J-S., Shires, J. and Siba Yahaya, M. (2024), *The role of the private sector in combatting gendered cyber harms: How private sector technologies can both endanger and advance gender-transformative cybersecurity*, Research Paper, London: Royal Institute of International Affairs, https://doi.org/10.55317/9781784136093.

This publication is printed on FSC-certified paper.
designbysoapbox.com

Independent thinking since 1920

CHATHAM HOUSE

The Royal Institute of International Affairs
Chatham House
10 St James's Square, London SW1Y 4LE
T +44 (0)20 7957 5700
contact@chathamhouse.org | chathamhouse.org

Charity Registration Number: 208223